

**Název práce:** Nové metody generování promluv v dialogových systémech

**Autor:** Ondřej Dušek

**Katedra:** Ústav formální a aplikované lingvistiky

**Vedoucí práce:** Ing. Mgr. Filip Jurčíček, Ph.D.,  
Ústav formální a aplikované lingvistiky

### **Abstrakt:**

Tato disertační zkoumá nové přístupy ke generování přirozeného jazyka (NLG) v hlasových dialogových systémech, tj. generování odpovědí systému pro uživatele. Zaměřuje se přitom na zlepšení adaptivity NLG ve třech ohledech: přenositelnost mezi různými doménami, přenositelnost mezi jazyky a přizpůsobení výstupu uživateli.

Ve všech ohledech dosahují naše generátory zlepšení oproti dřívějším přístupům: 1) Naše generátory, založené na statistických metodách (prohledávání  $A^*$  s perceptronovým rerankerem a architektuře rekurentních neuronových sítí sequence-to-sequence), lze natrénovat na datech bez podrobného sémantického zarovnání slov na atributy vstupní reprezentace, což dovoluje jednodušší přetrénování pro nové domény než předchozí přístupy. 2) Generátor založený na neuronových sítích dále rozšiřujeme tak, že při generování bere v potaz kontext dosavadního dialogu (tj. i uživatelův způsob vyjadřování) a vytváří tak výstup přizpůsobený uživateli. 3) Vyhodnocujeme také několik úprav systému založeného na neuronových sítích, které jsou zaměřeny na generování výstupu v morfologicky bohatých jazycích, a ukazujeme zlepšení v generování češtiny.

Při našich experimentech navíc porovnáváme různé architektury NLG (tradiciční dvojfázové zpracování s odděleným větným plánovačem a povrchovým realizátorem a integrovaný, jednofázový přístup). Pro trénování generátorů jsme též sestavili a zveřejnili dvě nové datové sady.

**Klíčová slova:** generování přirozeného jazyka, dialogové systémy, adaptivita, entrainment v dialogu, vícejazyčnost